

Continuous Dynamical Modes in Straits Having Arbitrary Cross Sections, with Applications to the Bab al Mandab*

LARRY J. PRATT AND HEATHER E. DEESE

Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

STEPHEN P. MURRAY

Coastal Studies Institute, Louisiana State University, Baton Rouge, Louisiana

WILLIAM JOHNS

Rosenthal School of Marine Science, University of Miami, Miami, Florida

(Manuscript received 14 June 1999, in final form 22 November 1999)

ABSTRACT

The continuous dynamical modes of the exchange flow in the Bab al Mandab are computed in an attempt to assess the hydraulic character of the flow at the sill. First, an extended version of the Taylor–Goldstein equation for long waves that accounts for cross-channel topographic variations, is developed. A series of calculations using idealized background velocity $U(z)$ and buoyancy frequency $N(z)$ are presented to illustrate the effects of simple topographic cross sections on the internal modes and their speeds. Next, hydrographic and direct velocity measurements from April to November 1996 using moored CTDs and a bottom-mounted ADCP are utilized to construct monthly mean vertical profiles of $N^2(z)$ and at the $U(z)$ sill. An analytical approximation of the true topography across the strait is also constructed. The observed monthly mean profiles are then used to solve for the phase speeds of the first and second internal modes. Additional calculations are carried out using a selection of “instantaneous” (2-h average) profiles measured during extremes of the semidiurnal tide. The results are compared with a three-layer analysis of data from the previous year.

Many of the authors’ conclusions follow from an intriguing observation concerning the long-wave phase speeds. Specifically, it was nearly always observed that the calculated speeds c_{-1} and c_1 of the two waves belonging to the first internal mode obey $c_{-1} < U_{\min} < U_{\max} < c_1$, where U_{\min} and U_{\max} are the minimum and maximum of the velocity profile. An immediate consequence is that neither wave has a critical level. For monthly mean profiles, each of which have $U_{\min} < 0 < U_{\max}$, the flow is therefore subcritical (the phase speeds of the two waves have opposite signs). For instantaneous profiles this relationship continues to hold, although the velocity profile can be unidirectional. Thus the flow can be critical ($c_{-1} = 0$ and/or $c_1 = 0$) or even supercritical (c_{-1} and c_1 have the same sign) with respect to the first mode. Similar findings follow for the second baroclinic mode phase speeds (c_{-2} and c_2). The authors conclude that hydraulically critical flow is an intermittent feature, influenced to a great extent by the tides. It is noted that the phase speed pairs for each mode lie very close to U_{\min} and U_{\max} . As suggested by the analysis of idealized profiles, this behavior is characteristic of flows that are marginally stable, perhaps as a result of prior mixing. This suggestion is supported by Richardson number (Ri) profiles calculated from the monthly mean and instantaneous data. Middepth values of Ri were frequently found to be $O(1)$ and sometimes $< 1/4$, a result consistent with the presence of mixing over portions of the water column.

1. Introduction

The general characteristics of the thermohaline circulation in the Red Sea and neighboring Gulf of Aden strongly suggest that the exchange flow in the con-

necting strait, the Bab al Mandab (BAM, Fig. 1), is hydraulically controlled. The Red Sea thermocline is very shallow, lying mainly above the (~ 165 m) sill depth (Maillard and Soliman 1986; Garrett et al. 1995) and the outflowing Red Sea water spills though the BAM and down into the Gulf of Aden (Fedorov and Mechanov 1988; Bower et al. 2000) in a manner reminiscent of flow spilling over a dam. The importance of understanding this presumed control (its location, the wave mode or modes involved, the locations of hydraulic jumps and other mixing agents, etc.) has been mentioned in connection with a variety of recent work on the Red Sea (Maxworthy 1997; Tragou and Garrett

* Woods Hole Oceanographic Institution Contribution Number 10067.

Corresponding author address: Dr. Larry J. Pratt, Woods Hole Oceanographic Institution, Dept. of Physical Oceanography, Woods Hole, MA 02543.
E-mail: lpratt@whoi.edu

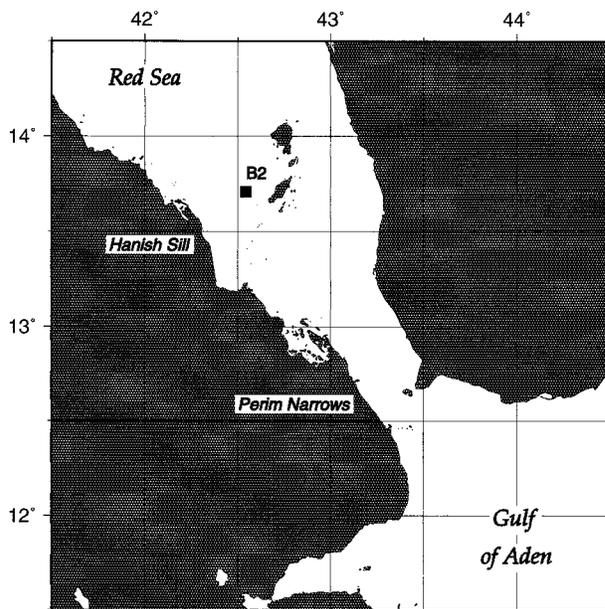


FIG. 1. The Bab al Mandab: **B2** marks the position of the ADCP and *T-S* mooring.

1997; Tragou et al. 1999; and especially Maxworthy 1997 and Smeed 1997, 2000).

Pratt et al. (1999, hereafter referred to as PJMK) recently assessed the internal hydraulics of the flow in the BAM by fitting a three-layer idealization of the exchange flow to density and velocity data. The reader is referred to the introduction of that paper and to Smeed 1997 for a more thorough discussion of the geometry of the strait and the seasonal character of the exchange flow. The data was collected (Murray and Johns 1997) as part of the first deployment of an April 1995–November 1996 field program employing CTD moorings and bottom mounted ADCPs, deployed at the Hanish Sill section and Perim Narrows section of the strait (Fig. 1). The three-layer model was used to calculate the long-wave propagation speeds of first and second internal modes at the sill and narrows. Surprisingly, the results indicated that the flow was substantially subcritical with respect to the first internal mode at the narrows, the location of strongest exchange velocities. As it turns out, the subcritical character of the flow is due to the large layer depths at the narrows, a feature that tends to enhance the speed of internal waves relative to the background flow. At the sill, where the velocities are weaker but the layer depths shallower, the nonsummer flow was found to be marginally subcritical with respect to the first mode, possibly indicating a section of critical flow (and hydraulic control) nearby. Another surprise finding was that the flow during nonsummer months was nearly critical with respect to the second internal mode at the sill (and possibly the narrows). The vertical structure of this mode suggests that the corresponding hydraulic control might regulate the composition of the

Red Sea outflow, particularly the relative contributions of upper Red Sea deep water and Red Sea intermediate water.

The study of PJMK raised a number of issues and was subject to several limitations, many of which are addressed in the present study. First of all, the velocity and density structure in the BAM does not have a distinct layer character. The velocity varies continuously and the density, while relatively uniform near the top and bottom of the water column, varies continuously and smoothly in the interior. There are no obvious interfaces and PJMK were forced to vary the positions of their model interfaces over what was believed to be reasonable bounds in order to make uncertainty estimates in the computed phase speeds. Second, there were a number of gaps in the hydrographic data that were filled from extrapolation or historical sources. Finally, all calculations were based on monthly mean vertical profiles, so little information about instantaneous flows and the effects of tides and other sources of time-dependence was found.

Since PJMK was published, data from the second deployment (April through November 1996) has become available. Although this dataset is limited by the fact that no hydrographic measurements were taken at the narrows, the measurements at the sill are more complete than those from the first deployment. (This is fortuitous since the results of PJMK seem to indicate the sill as the most likely location of hydraulic control with respect to the first mode.) The completeness of the velocity and hydrographic measurements at the sill allows us to compute the first and second continuous dynamical modes and their phase speeds. In order to do so we first extend the theory for calculating long-wave modes to include the possibility of variable cross-channel bottom topography. We suspect that topography is important in the BAM because the outflowing Red Sea Water is confined to a narrow deep channel running along the center of the strait, as shown in Fig. 2a. The computed phase speeds allow us to make determinations regarding the hydraulic character (subcritical, supercritical, or critical with respect to different modes) of the flow. We do this using both monthly mean profiles and a selection of “instantaneous” profiles. The latter are really 2-h average profiles that coincide with maximum tidal inflow and outflow and with different states of net (barotropic) inflow/outflow due to longer period disturbances. The calculated phase speeds may imply vertical mixing, and we discuss this in connection with calculations of the Richardson number over the water column.

The connection between long waves and hydraulic behavior is well understood in layered models, and the reader is referred to Smeed (2000) and PJMK for discussion of this issue in relation to the BAM. In continuously stratified flows, which exhibit more complex behavior and are more difficult to model, this connection is not as well established. Nevertheless, it is clear that signals propagating from the Indian Ocean into the Red

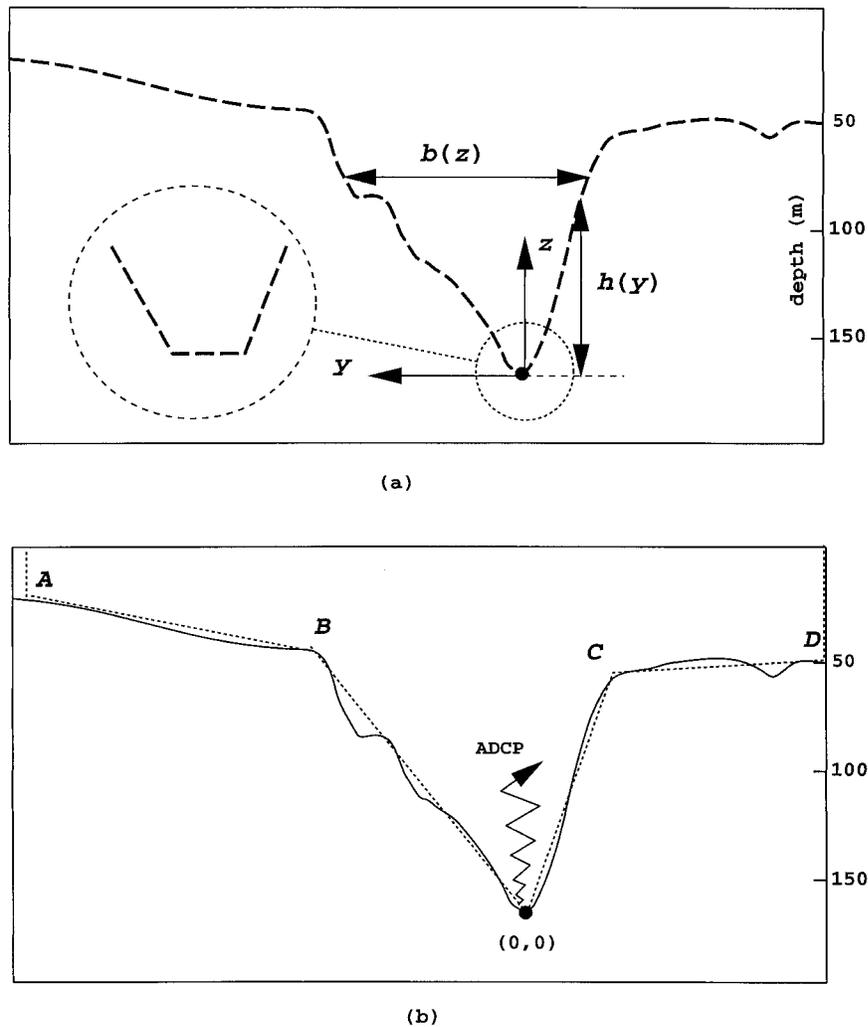


FIG. 2. (a) Bottom topography across the Hanish Sill section. The observer faces toward the Red Sea. The insert shows a local, flat-bottom approximation to the topography at the deepest point that can be used to prevent the function $T(z)$ from becoming singular. (b) The actual topography (solid line) with a piecewise-linear approximation (dashed line). The end points of the dashed line segments are given by $A = (-23 \text{ km}, 150 \text{ m})$, $B = (-10 \text{ km}, 125 \text{ m})$, $C = (4 \text{ km}, 110 \text{ m})$, and $D = (14 \text{ km}, 120 \text{ m})$, all referenced to the deepest point of the cross section. At the sill, the bottom is approximated by a horizontal section 100 m long. The channel terminates in vertical side walls when the depth becomes less than 10 m (not shown) and this results in a total width of approximately 50 km.

Sea potentially influence features like the structure of the Red Sea thermocline, the composition of the outflow, and perhaps even elements of the horizontal circulation. It is therefore important to assess the ability of these signals to reach the Red Sea. If certain types of signals are unable travel northwest through the BAM, then the Red Sea possesses a degree of isolation and self-determination that goes hand-in-hand with hydraulic control. If all signals can travel in this direction, than the BAM flow would not be hydraulically controlled, and one might expect features like the depth of the Red Sea thermocline to be imposed largely by the vertical density structure of the Gulf of Aden and Indian Ocean. In brief,

there is sufficient motivation for calculating the phase speeds of these signals at the sill even though one may not have a particular model of hydraulic behavior for a continuously stratified exchange flow in hand.

A number of calculations and results are new and distinguish this work from PJMK. First, we introduce a method for calculating the continuous, dynamical, internal long-wave modes in the presence of arbitrary, cross-channel topography by deriving an extended version of the Taylor–Goldstein equation for long waves. Second, we employ a few idealized profiles of channel topography, background velocity and N^2 to illustrate some of the basic features of long-wave modes expected

in exchange flows over typical topography. Third, we document certain observed features, including monthly mean and selected 2-h average velocity, stratification, and Richardson number for the second deployment (April–November 1996). (PJKM used data from the first deployment, which spanned June 1995–March 1996). Fourth, we compute the lowest two dynamical modes, their phase speeds, and the corresponding uncertainty bars for the second deployment monthly mean sill flows and use the results to assess the criticality of the exchange flow. (In order to overlap the first deployment and develop a full year of results, we also perform the calculations for a block of first-deployment months.) Finally, to gain some insight into the effects of time dependence, we perform the same calculations for a selection of instantaneous flows.

2. An extended Taylor–Goldstein equation for long waves

The BAM outflow is largely confined to a deep central trough whose width is much less than the width of the strait (PJKM). This topographic feature can be seen in the Hanish Sill cross section shown in Fig. 2a. We formally analyze the general effects of cross-channel topographic variations by considering the propagation of long, internal gravity waves in a channel with laterally varying but longitudinally uniform bottom elevation. The waves are linear and propagate in a stratified background flow with velocity $U(z)$ and density $\rho_0(z)$. The channel is aligned in the x direction and the bottom elevation $h(y)$ has a single minimum with respect to the cross-channel coordinate y . As shown in Fig. 2a, the width of the channel at any elevation z is denoted $b(z)$. If several minima in $h(y)$ are present, then $b(z)$ is understood to represent the sum of the widths of the individual topographic troughs.

Let ρ , u , and p denote small perturbations from the density, x velocity, and hydrostatic pressure of the background flow, and let v and w denote the associated lateral and vertical velocity components. Employing the Boussinesq approximation, the linearized, inviscid, hydrostatic equations of motion describing these fields are then given by

$$\rho_0 \left[\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) u + \frac{dU}{dz} w \right] = - \frac{\partial p}{\partial x} \quad (2.1)$$

$$\rho_0 \left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) v = - \frac{\partial p}{\partial y} \quad (2.2)$$

$$\frac{\partial p}{\partial z} = - \rho g \quad (2.3)$$

$$\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \rho + \frac{d\rho_0}{dz} w = 0 \quad (2.4)$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0. \quad (2.5)$$

We seek waves for which ρ , p , u , and w are uniform in y , implying that the isopycnal surfaces rise and fall uniformly across the channel. Such solutions are dynamically consistent only in the limit of long wave length compared to channel width.¹ Integrating (2.5) across the channel at any z and applying the conditions $w = v(dh/dy)$ at the two side walls leads to

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} + T(z)w = 0, \quad (2.6)$$

where $T(z) = b^{-1}db/dz = d(\ln b)/dz$.

Now let $(u, w, \rho, p) = \text{Re}[(\tilde{u}(z), \tilde{w}(z), \tilde{\rho}(z), \tilde{p}(z))e^{ik(x-ct)}]$, so that (2.1), (2.3), (2.4), and (2.6) become

$$\rho_0(U - c)\tilde{u} + \frac{\rho_0}{ik} \frac{dU}{dz} \tilde{w} = -\tilde{p} \quad (2.7)$$

$$\frac{d\tilde{p}}{dz} = -\tilde{\rho}g \quad (2.8)$$

$$(U - c)\tilde{p} + \frac{\tilde{w}}{ik} \frac{d\rho_0}{dz} = 0 \quad (2.9)$$

$$ik\tilde{u} + \frac{\partial \tilde{w}}{\partial z} + T\tilde{w} = 0. \quad (2.10)$$

Eliminating all variables in favor of \tilde{w} and neglecting derivatives of $\rho_0(z)$ unless multiplied by g leads to

$$(U - c) \frac{d^2 \tilde{w}}{dz^2} + \left[\frac{N^2}{U - c} - \frac{d^2 U}{dz^2} \right] \tilde{w} + \frac{d}{dz} [(U - c)T\tilde{w}] = 0, \quad (2.11)$$

where $N^2 = -\rho_0^{-1}g \, d\rho_0/dz$.

The geometry of the cross section is contained in the parameter $T(z)$. If the cross section is rectangular ($T = 0$), the Taylor–Goldstein equation for long waves is recovered. For the calculations presented later, we will assume a rigid-lid boundary condition at the upper surface ($z = z_T$) and thus $\tilde{w}(0) = \tilde{w}(D) = 0$, where $z = 0$ denotes the elevation of the deepest point in the channel and $z = D$ denotes the upper surface elevation.

In further analysis it will be convenient to nondimensionalize (2.11) by scaling U and c by an unspecified background velocity scale U_0 , \tilde{w} by an unspecified vertical velocity scale W , and z , h , and y by D . (Thus, dimensional N^2 is replaced by $N^2 D^2 / U_0^2$ and dimensional T by DT .) The dimensionless problem for the vertical structure of the modes then becomes

¹ If the wavelength λ is large compared to the channel width b , then the continuity equation suggests that $v/u = O(b/\lambda)$. The momentum equations (2.1) and (2.2) then suggest that the along-channel pressure gradient is larger than the cross-channel pressure gradient by the same ratio. Under this condition, it not difficult to show from (2.1)–(2.5) that cross-channel gradients in the dependent variables remain negligible if they are initially zero.

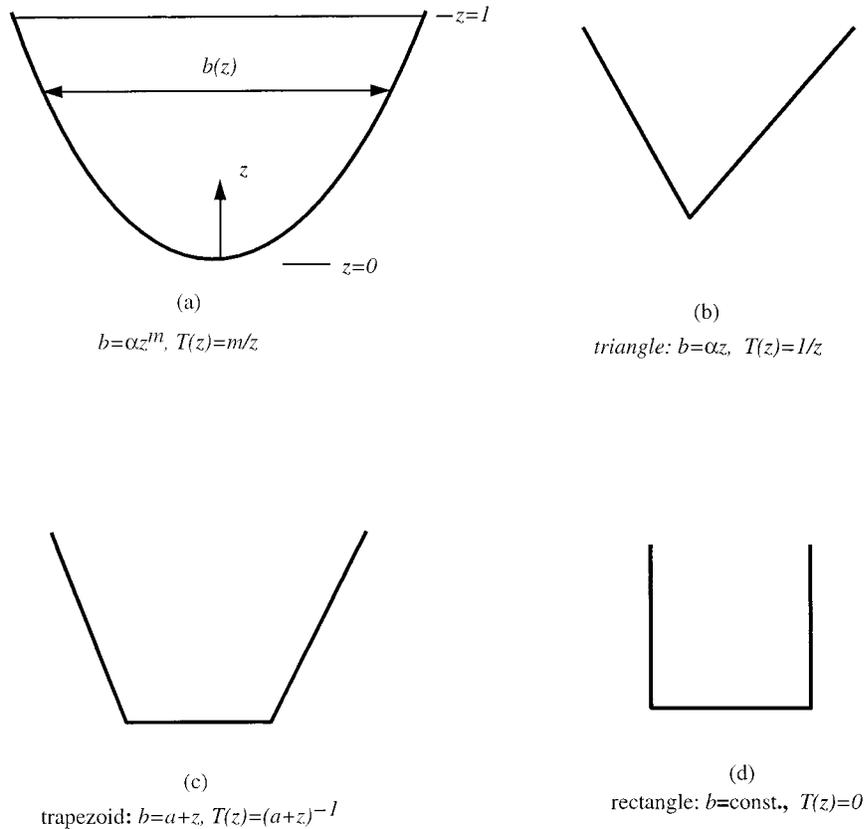


FIG. 3. Various idealized topographic profiles along with their nondimensional $T(z)$ [$=D$ times dimensional $T(z)$].

$$\frac{d^2 \tilde{w}}{dz^2} + \left[\frac{N^2}{(U - c)^2} - \frac{1}{(U - c)} \frac{d^2 U}{dz^2} \right] \tilde{w} + \frac{1}{(U - c)} \frac{d}{dz} [(U - c) T \tilde{w}] = 0 \quad (2.12)$$

with

$$\tilde{w}(0) = \tilde{w}(1) = 0. \quad (2.13)$$

In case studies of geophysical interest, where b smoothly goes to zero as $z \rightarrow 0$, $T(z)$ typically has a z^{-1} singularity. For example, let $h(y) \propto \hat{\alpha} y^\mu$ near $y = 0$, so that $b = \alpha z^m$, where $m = \mu^{-1}$ and $\alpha = 2\hat{\alpha}^{-1/\mu}$. It follows that $T(z) = (d/dz) \ln(\alpha z^m) = m/z$. Figure 3 shows some idealized shapes with corresponding $T(z)$. The topographic shapes shown in Figs. 3c,d have finite b at $z = 0$ and thus avoid the singularity. For numerical purposes, the singularity can be avoided by approximating the actual bottom topography at $z = 0$ by a short horizontal section, as suggested in the inset of Fig. 2a. This subject discussed further in appendix A.

The broken dashed line in Fig. 2b shows an approximation to the Hanish Sill topography (solid curve) used in our calculations. The deepest part is represented by a 100-m long horizontal section at 160-m depth. The slopes of the side walls are varied in a piecewise-con-

stant manner in order to fit the actual topography. Above a level z_0 the side walls become vertical ($T = 0$). The dimensionless shape function describing the approximate topography is

$$T(z) = \begin{cases} (z + 0.0193)^{-1} & (0 \leq z < 0.668) \\ (z - 0.5947)^{-1} & (0.668 < z < 0.762) \\ (z - 0.6440)^{-1} & (0.762 < z < z_0) \\ 0 & (z_0 < z \leq 1), \end{cases} \quad (2.14)$$

where $z_0 = 15/16$.

3. Examples of modes in idealized systems

Since our assessment of the hydraulics of the BAM sill flow is centered on signal propagation, it is important to understand in advance what types of signals are possible and what restrictions exist on their phase speeds. We therefore begin this section by reviewing the properties of solutions to the original Taylor–Goldstein equation. Unless noted otherwise, these results apply directly to our extended, long-wave version of Eq. (2.12) provided that the singularity in $T(z)$ at $z = 0$ is no worse than $1/z$. That is, the cross-sectional geometry of the

channel has not been found to alter basic theorems about phase speed bounds and the Richardson number criterion for instability. Proofs showing the extension of previous results [such as Howard's (1961) semicircle theorem] to the new equation have been developed by J. Deng, L. Howard, L. Pratt, H. Deese and C. K. R. T. Jones (unpublished manuscript) but these are beyond the scope of the present paper. The influence of topography on the structure and propagation of wave modes becomes evident through specific examples, and we present several later in this section.

As summarized by Drazin and Reid (1981) and Baines (1995) the solutions of the (original) Taylor–Goldstein equation with $N^2 > 0$ depend crucially on the minimum value of the Richardson number

$$\text{Ri} = \frac{N^2(z)}{(dU/dz)^2}$$

over the water column. If $(\text{Ri})_{\min} > 1/4$, then there is a countable infinity of discrete neutral modes. These can be thought of as internal gravity waves that have been modified by the background shear. Denoting the corresponding eigenfunctions and eigenvalues by $[\tilde{w}_j(z), c_j]$, it can be shown that

$$\begin{aligned} U_{\min} - \frac{N_{\max}}{(\pi^2 + k^2)^{1/2}} &< c_{-1} < c_{-2} < c_{-3} < \dots < U_{\min} \\ &< U_{\max} < \dots < c_3 < c_2 < c_1 \\ &< U_{\max} + \frac{N_{\max}}{(\pi^2 + k^2)^{1/2}}, \end{aligned} \quad (3.1)$$

and $c_j \rightarrow U_{\min}$ as $j \rightarrow -\infty$ and $c_j \rightarrow U_{\max}$ as $j \rightarrow \infty$. Thus the wave speeds lie outside the range of the basic velocity but within a finite factor ($\pi^{-1}N_{\max}$ for long waves) of the velocity extremes. Finally, it can also be shown that the j th eigenfunction $\tilde{w}_j(z)$ has exactly $|j| - 1$ zero crossings in $0 < z < 1$.

The discrete neutral modes are important as they provide a connection with layered systems. An n -layer hydraulic model has n normal modes and these are analogous to the first n discrete modes of a continuous system. This analogy can break down when $(\text{Ri})_{\min} < 1/4$ in which case the flow may be unstable (Miles 1961). Neutrally stable discrete modes can still exist but their number may be finite, even zero. Their phase speeds continue to lie outside the range of the basic velocity. In addition, there may be a finite number of nonsingular unstable modes (with conjugate damped partners). According to the semicircle theorem, the real part of the phase speed for such a mode will lie within the range of the background velocity. Finally there can be a finite number of marginally stable² neutral modes that have real phase speeds c in the range of the background ve-

TABLE 1. Phase speed eigenvalues for three idealized topographies.

Case	$ c_{\pm 1} /N$	$ c_{\pm 2} /N$	$ c_{\pm 3} /N$
$T(z) = 0$ (Rectangular bottom)	0.318	0.159	0.106
$T(z) = 1/(z + \alpha)$ (Trapezoid with $\alpha = 1$)	0.312	0.158	0.105
$T(z) = 1/z$ (Triangular bottom)	0.261	0.143	0.103

locity and therefore have critical levels z_c such that $U(z_c) = c$. If the value of the Richardson number at the critical level is $> 1/4$, $\tilde{w}(z)$ oscillates rapidly in the neighborhood of z_c (Booker and Bretherton 1967).

In addition to the discrete modes, a continuous spectrum of neutral modes will exist with critical levels. At these critical levels, $\tilde{w}(z)$ will have a singularity no worse than a discontinuity in the first derivative. To the best of our knowledge, the role that these waves might play in the hydraulic behavior of the flow is unknown. However, one can easily examine the behavior of Couette flow (which has a continuous spectrum but no discrete modes) in the presence of an obstacle and show that there can be no hydraulic control. This calculation is presented in appendix B.

Now consider some examples of long-wave modes in idealized systems, starting with the case $U(z) = 0$ and $N = \text{const}$. Here the eigenvalues arise in pairs, designated $j = \pm 1, \pm 2$, etc., with $c_j = -c_{-j}$. If the channel cross-section is rectangular ($T = 0$) then $c_j = \pm N/j\pi$ and $\tilde{w}_j(z) = a_j \sin(j\pi z)$. The vertical structure function $\tilde{w}_j(z)$ is the same for each of the waves having the same $|j|$, but the phase relation between $\tilde{w}_j(z)$ and the pressure $\tilde{p}_j(z)$ differ for each pair. When a background velocity is present and $(\text{Ri})_{\min} > 1/4$, it is still possible to group the waves in pairs. Although $\tilde{w}_j(z)$ generally differs from $\tilde{w}_{-j}(z)$, the number of zero crossings is the same for each. Thus, we will consider each $|j|$ as designating a vertical mode, recognizing that each such mode is associated with two waves.

Next, consider the topographic function $T(z) = 1/(z + \alpha)$, $\alpha = y_0/(y_1 - y_0)$, corresponding to a trapezoid with bottom and top dimensions y_0 and y_1 . A triangle results from the limit $\alpha \rightarrow 0$, a rectangle from $\alpha \rightarrow \infty$. Solutions to (2.12) with $U = 0$ and $N^2 = \text{const}$ take the form

$$\tilde{w}_j(z) = C_{1j} J_1 \left(\frac{N(z + \alpha)}{|c_j|} \right) + C_{2j} Y_1 \left(\frac{N(z + \alpha)}{|c_j|} \right), \quad (3.2)$$

where J_1 denotes the Bessel function of first kind and Y_1 denotes Weber's Bessel function of the second kind, both of order 1. Application of the boundary conditions (2.13) yields a transcendental relation for the eigenvalues c_j , and Table 1 summarizes these speeds for rectangular, trapezoidal (with $\alpha = 1$), and triangular topography. As one moves from the case of a rectangular channel (for which $c_j = \pm N/j\pi$) to a trapezoid, and eventually to a triangle, the phase speed for each mode decreases. This trend is not surprising: the average depth of the fluid

² That is, contiguous to an unstable mode.

decreases in each case. There is also a tendency for the wave modes, particularly the higher ones, in trapezoidal or triangular geometry, to be intensified near the bottom, as evidenced by the decay of the Bessel functions for large arguments: $J_1(z) \rightarrow \sqrt{2/(\pi z)} \cos(z - 3\pi/4)$ and $Y_1(z) \rightarrow \sqrt{2/(\pi z)} \sin(z - 3\pi/4)$ as $z \rightarrow \infty$.

Some insight into the influence of a background flow can be gained by considering a linear shear flow $U(z) = U_{\max}(2z - 1)$ with $N^2 = \text{const}$, contained in a rectangular channel ($T = 0$). The value of Ri is uniform and for $\text{Ri} < 1/4$ the flow is stable and has no discrete modes (Case 1960). For $\text{Ri} > 1/4$ the discrete modes are given (Baines 1995) by

$$\tilde{w}_j(z) = \left[1 + \frac{U_{\max}(1 - 2z)}{c_j} \right]^{1/2} \times \sin \left\{ \mu \ln \left[\frac{1 + \frac{U_{\max}(1 - 2z)}{c_j}}{1 + \frac{U_{\max}}{c_j}} \right] \right\},$$

where

$$c_j = \coth(j\pi/2\mu), \quad j = \pm 1, \pm 2, \dots$$

and where $\mu = (\text{Ri} - 1/4)^{1/2}$. The phase speeds for $|j| = 1, 2, 3$, and 4 are plotted as a function of Ri in Fig. 4a. As $\text{Ri} \rightarrow 1/4$ from above, the positive phase speeds converge to the value U_{\max} (and the negative phase speeds to $-U_{\max}$) so that the propagation speeds accumulate around the upper and lower limits of the velocity range. This behavior is quite important in interpreting later results for the BAM data.

Figure 4b shows $\tilde{w}_j(z)$ for $j = 1, 2$, and 3 at values $\text{Ri} = 4.25$ and 16.25. Decreasing the value of Ri causes the zero crossings to migrate upward in the water column so that the wiggles become confined near the top boundary. (For the negative j modes, the wiggles would become densely packed at the bottom of the water column.) The rapid oscillations are associated with the emergence of a critical level at $z = 1$ (or $z = 0$), which occurs for each mode as $\text{Ri} \rightarrow 1/4$. This behavior, which was alluded to above, clearly raises an important issue concerning resolution of the higher modes. As Ri is reduced toward $1/4$, the eigenfunctions for different j look alike over most of the depth range (as shown by the upper trio of solutions in Fig. 4b). The differences can be distinguished only in a narrow band at the top (or bottom) of the water column.

A final example combining topographic elements with shear effects arises from consideration of

$$U(z) = U_{\max} \frac{\tanh(4z - 2)}{\tanh(2)}, \quad (3.3a)$$

$$T(z) = 1/z \quad (3.3b)$$

with $N^2 = \text{const}$. This background state consists of an

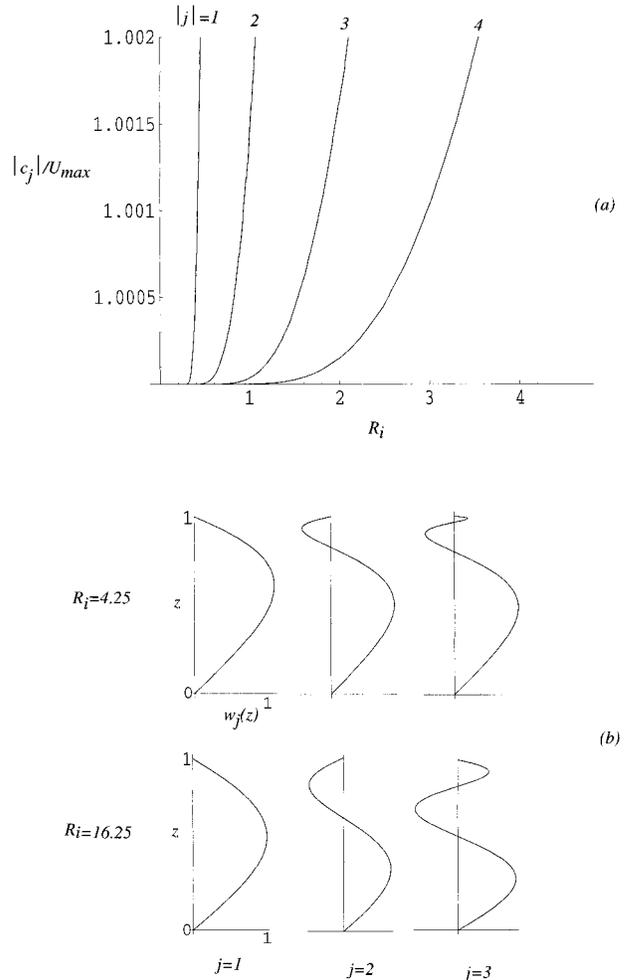


FIG. 4. (a) The phase speeds for modes $|j| = 1, 2, 3$, and 4 of the background state $U(z) = U_{\max}(2z - 1)$, $N^2 = \text{const}$, and rectangular topography $T(z) = 0$. The speeds are plotted against the Richardson number (which is constant with z). Note that $c_j = -c_{-j}$. (b) The vertical structure functions for modes $|j| = 1, 2, 3$ for $\text{Ri} = 4.25$ and $\text{Ri} = 16.25$.

exchange flow in a channel with a triangular cross section. The phase speeds of the first few modes have been calculated numerically and the results are presented in Figs. 5a,b. As the minimum value of Ri (which here occurs at middepth) is reduced from large to $O(1)$ values, the phase speeds of the modes converge to $\pm U_{\max}$ in a manner similar to those of the case just described. Figure 5a shows this behavior for the phase speeds of the wave pairs for the ($j = \pm 1, \pm 2$, and ± 3). Unlike the constant shear, rectangular geometry case, however, these modes do not disappear when $(\text{Ri})_{\min} = 1/4$ but, instead, continue to $(\text{Ri})_{\min} = 0$. The vertical structure of the modes (not shown) is similar to what was observed in the previous example with the zero crossings migrating toward the level of U_{\min} and U_{\max} , here the bottom and top of the channel, as $(\text{Ri})_{\min} \rightarrow 0$.

The convergence of c_j and c_{-j} just above and below

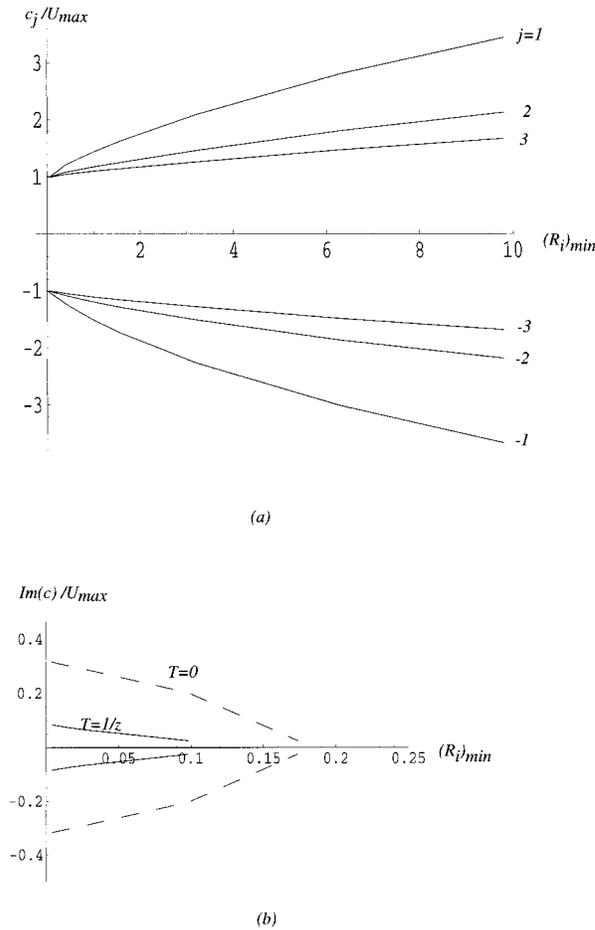


FIG. 5. Phase speed (a) normalized by the maximum velocity U_{max} , for modes $|j| = 1, 2, 3$ of the velocity profile $U(z) = U_{max} \tanh(4z - 2)/\tanh(2)$, with triangular topography ($T(z) = 1/z$), and $N^2 = \text{const}$. The results are plotted against the minimum Richardson number, which occurs at middepth. (b) Enlarged portion of (a) near the origin and showing $Im(c)/U_{max}$ for the unstable mode and its damped partner (solid lines). This mode has $Re(c)/U_{max} \cong 0.40$ over the range shown, which gives a critical level just below middepth. The dashed lines give $Im(c)/U_{max}$ for the same mode in the presence of rectangular bottom topography [$T(z) = 0$]. This mode has a critical level at $z = 0.5$.

U_{min} and U_{max} as the stability of the flow is reduced can be anticipated by writing (3.1) in a slightly different form. Consider a velocity profile of fixed shape in that U_{max} can be increased or decreased but remains proportional to U_{min} . Dividing (3.1) by U_{max} or U_{min} leads to the dimensionless bounds

$$1 < \dots < \frac{c_3}{U_{max}} < \frac{c_2}{U_{max}} < \frac{c_1}{U_{max}} < 1 + \frac{R_{bulk}^{1/2}}{\pi}$$

and

$$1 - \frac{U_{max}}{\pi U_{min}} R_{bulk}^{1/2} < \frac{c_{-1}}{U_{min}} < \frac{c_{-2}}{U_{min}} < \frac{c_{-3}}{U_{min}} < \dots < 1,$$

where R has been set to zero and

$$R_{bulk} = \left(\frac{N_{max}}{U_{max}/D} \right)^2.$$

As the overall Richardson number R_{bulk} is reduced, all of the c_j approach U_{max} from above and all of the c_{-j} approach U_{min} from below. (Of course, some or all of the discrete neutral modes may cease to exist as R_{bulk} is reduced.)

The background state (3.3) also has an unstable normal mode, and we briefly note the effect of topographic variations on this instability. For a rectangular cross section [$T(z) = 0$] a mode with complex c first appears at $(Ri)_{min} \cong 0.17$. The upper dashed curve of Fig. 5b shows $Im(c)$ as a function of $(Ri)_{min}$ and reveals a maximum $Im(c) \cong 0.32U_{max}$ at $(Ri)_{min} \cong 0$. With the triangular topography, the unstable mode appears at a lower $(Ri)_{min}$ ($\cong 0.09$) and has a smaller maximum ($\cong 0.09U_{max}$). The implication is that the topography $T(z) = 1/z$ has a stabilizing influence on long waves, although this has not been tested in connection with other velocity profiles.

This section contains quite a bit of information, and we now reemphasize some of the behavior that could be of importance in the interpretation of results from the BAM or other straits. The first concerns the general effect of “realistic” topography, meaning any cross section containing significant decreases in width with depth. With $U = 0$ and $N^2 = \text{const}$, realistic topography causes a moderate reduction in c below the case of a rectangle (Table 1). The topographic influence on c in the presence of nontrivial $U(z)$ is more difficult to describe simply. In any case, one should proceed under the assumption that topography will be at least moderately important in most straits of geophysical interest. Second, propagating modes in the presence of topography are subject to the same phase speed bounds that govern modes in a rectangular channel. For instance, the speeds of discrete neutral modes must lie outside the extremes U_{min} and U_{max} of the velocity profile. As the bulk Richardson number becomes small, these speeds accumulate just outside of U_{min} and U_{max} . Under these conditions each mode nearly has a critical level. An important consequence of the phase speed bounds for an exchange flow ($U_{max} > 0$ and $U_{min} < 0$) is that all c_j are nonzero. Therefore, such a flow cannot be hydraulically critical ($c_j = 0$) with respect to any discrete neutral mode. If the velocity profile becomes temporarily unidirectional through the action of a barotropic tide, say, then hydraulic criticality is possible. The occurrence of phase speeds just outside the range of the basic velocity and the occurrence of critical flow only when the velocity profile is unidirectional is exactly the situation we observe in the BAM.

4. Monthly mean stratification, velocity, and Richardson number

As part of the second deployment, temperature and conductivity were measured using five Seacats moored

TABLE 2. Coefficients for σ_θ fit [Eq. (4.1)].

Month/ 1996	s_1	s_2	s_3	s_4
Apr	25.944	2.578	72.0	28.0
May	25.723	2.798	68.0	26.0
Jun	25.154	3.030	60.0	42.0
Jul	25.010	3.160	52.0	42.0
Sep	25.047	3.008	56.0	40.0
Oct	25.817	2.517	69.5	21.5
Nov	26.216	2.207	78.0	20.0

at depths of 23 m, 51 m, 62 m, 114 m, and 142 m at the position marked B2 in Fig. 1. We computed monthly mean values of σ_θ at each of these levels and then least square fitted the five values for months April–July and September–November 1996 to the curve

$$\sigma_\theta(z) = s_1 + s_2 \{ \tanh[(z - s_3)/s_4] \}. \quad (4.1)$$

The coefficients s_n for each month are listed in Table 2. For month 8, when the flow includes a significant surface outflow and the Gulf of Aden intrusion is lower in the water column (as discussed below), (4.1) does not sufficiently describe the flow; we instead utilized a piecewise linear fit. The fitted $\sigma_\theta(z)$ for months 4, 8,

and 11 are shown along with the associated dimensional N^2 in Fig. 6.

We calculated monthly mean ADCP velocity profiles with 10-m vertical resolution extending from 10-m to 130-m depth as shown in Fig. 7. We linearly extrapolated from neighboring points to obtain velocities at 0 m, 140 m, and 150 m (represented by crosses in Fig. 7). During the nonsummer months the exchange flow is characterized by moderate surface inflow (positive velocity) and significantly stronger deep outflow. As noted by PJMK, the inflow velocity is weaker because it is spread over a much larger cross-sectional area than the outflow, the latter being confined to the central topographic trough. In fact, the inflow and outflow volume transports nearly cancel, as required by the overall mass balance for the Red Sea.³ During the “summer” (months 6, 7, 8, and 9) the deep outflow is attenuated while the inflow lies lower in the water column. Only during months 8 and 9 is there a clear surface outflow.

The Richardson numbers for each month are dis-

³ Replenishment of water lost to evaporation requires a net inflow of only a few centimeters per second.

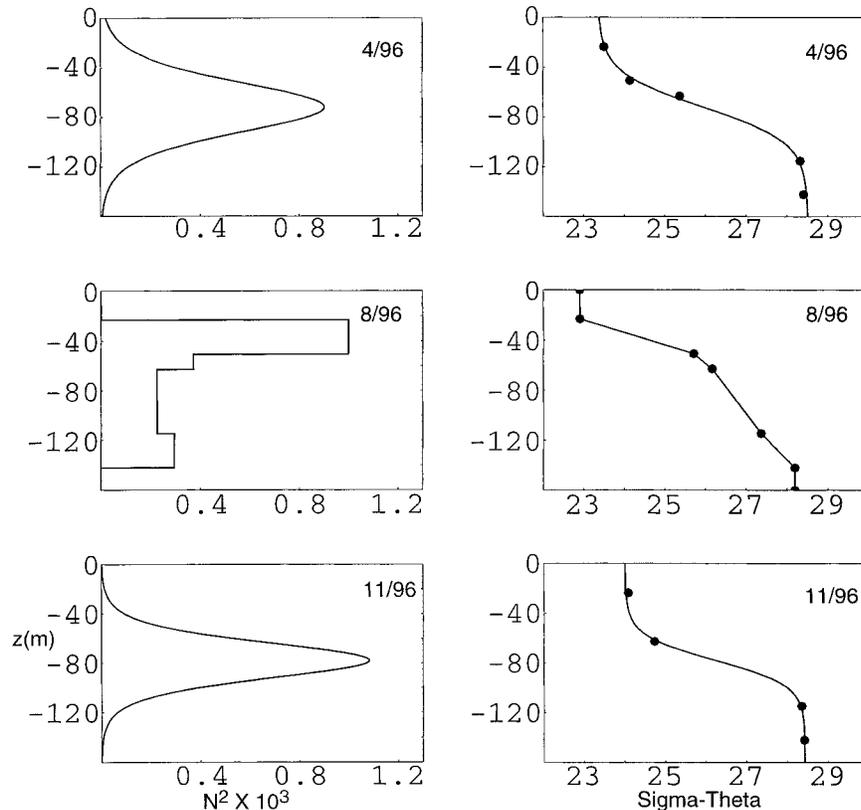


FIG. 6. The right frames show analytical fits to the monthly mean σ_θ and the left frames show the associated N^2 (in s^{-1}) for months 4, 8, and 11 of 1996. The analytical fit was obtained using (4.1) for each month but 8, where a piecewise linear fit was used. The measurements of σ_θ were by CTDs moored at the sill and are indicated by points in the right-hand frames.

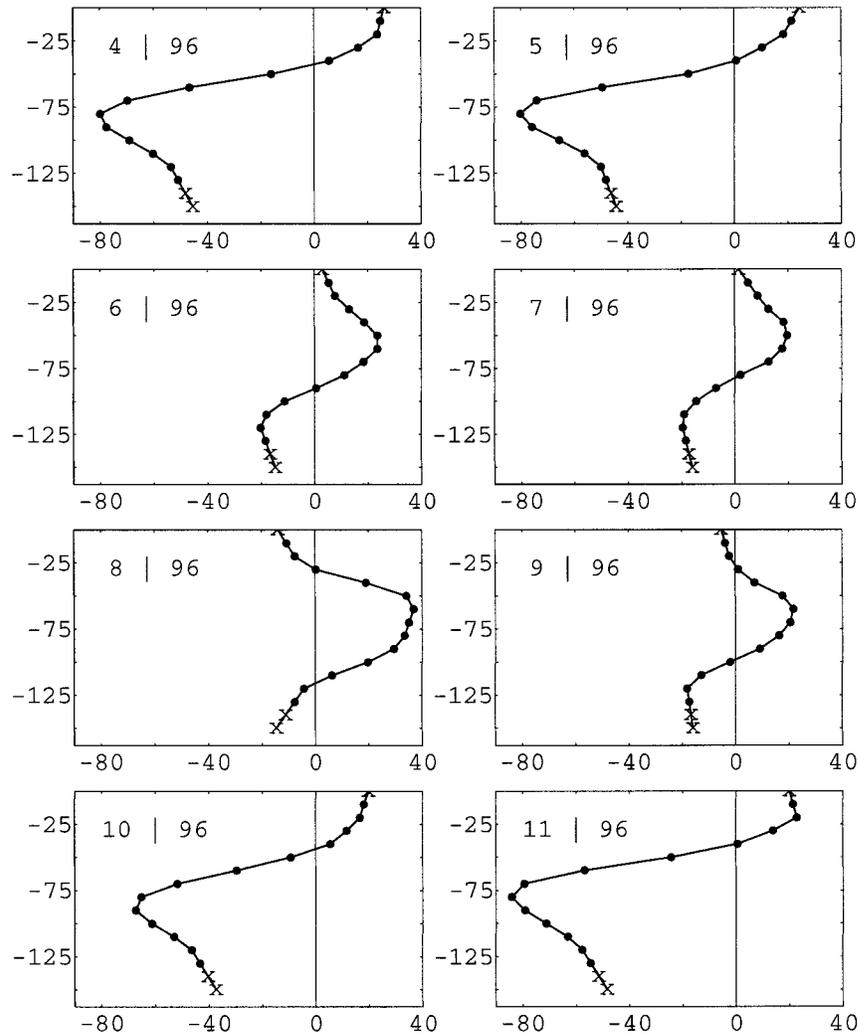


FIG. 7. Monthly mean ADCP sill velocities (cm s^{-1}) for months 4–11 of 1996. Dots indicate measurements averaged over 10-m bins while \times 's indicate values filled in using a linear extrapolation. The depth below surface is in meters.

played in Fig. 8 along with uncertainty bars.⁴ (At depths where the uncertainty is quite small, the central estimated value is represented by a dot.) In addition to analyzing the second deployment months (April–November 1996), we reappraise months November 1995–March 1996 of the first deployment. Since we do not have stratification data from the first deployment, we utilize $N^2(z)$ from both April 1996 and November 1996 in our calculations of Ri , resulting in larger uncertainty bars for these months.

⁴ Uncertainty bars reflect uncertainties due to the finite difference approximation of the shear, different extrapolation schemes used at the top and bottom of the water column, and alternative fits to the density data used for the month of August. See appendix C for details.

Although instability is formally possible only for

$$Ri < 1/4,$$

values of about one or less can be suggestive of prior mixing. Such values occur during all months except June, July, August and September 1996, all four of which are characterized by relatively weak flows. Comparison of Figs. 7 and 8 shows that the smallest Ri for a given month generally occur within 10–20 m of the mean interface (the zero crossing of mean velocity profile). The position of this mean interface between inflowing Gulf of Aden water and deeper outflowing Red Sea water ranges from 35 m in winter to 110 m in summer. (The latter roughly corresponds to the depth of the lower maxima in the August N^2 , as shown in the middle left panel of Fig. 6.) Although some very

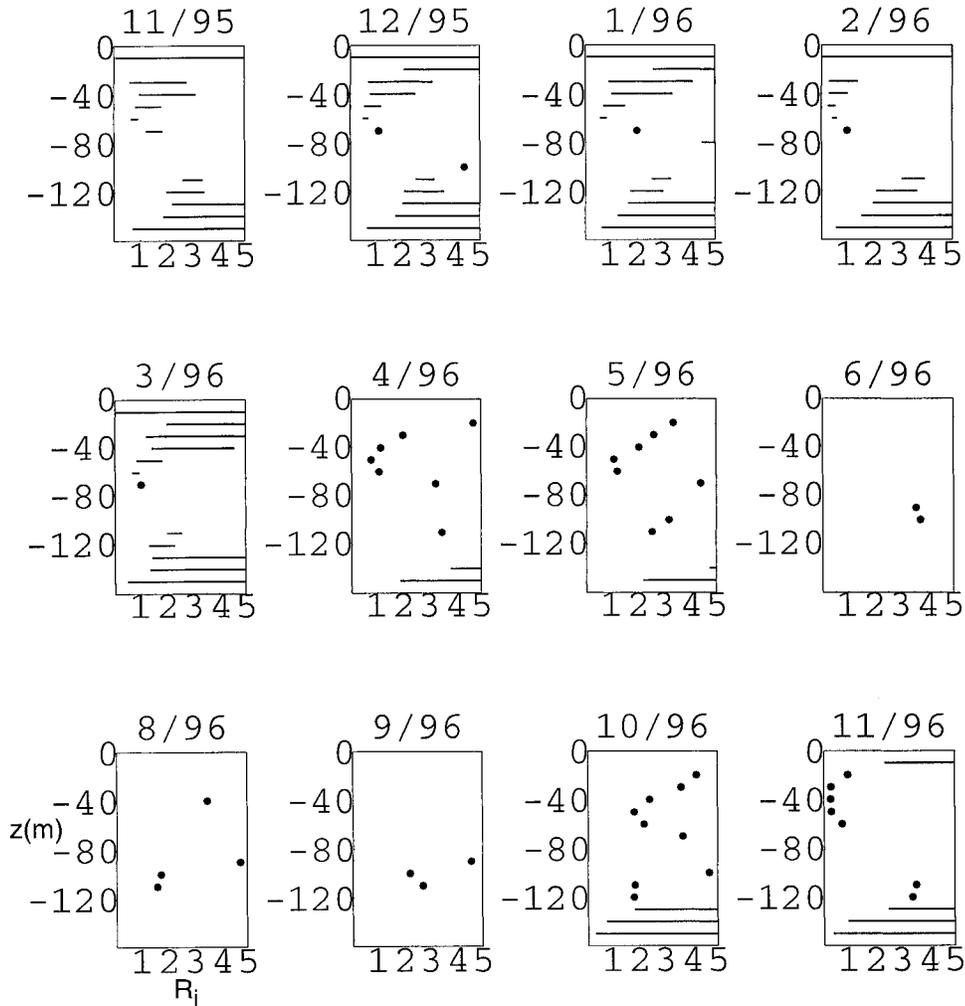


FIG. 8. Richardson number, $Ri = N^2/(dU/dz)^2$, for the monthly mean flows over Nov 1995–Nov 1996 (11/95–11/96). A horizontal bar at a particular depth gives the range in uncertainty in the estimated value of Ri at that depth due to interpolation or extrapolation of the velocity and N^2 measurements (see appendix C). Dots represent calculations for which the uncertainty bars are tiny. When no values appear, such as for most of the depths in month 6/96 the minimum Ri is > 5 . The results for 7/96 are not shown since $Ri > 5$ at all levels.

low estimates of Ri appear in the top 20 m and bottom 30 m of the water column in several months, the uncertainty is magnified at these depths because both the shear and stratification are weak and the velocity data is extrapolated (see appendix C).

5. Criticality of the monthly mean sill flow with respect to the first and second internal modes

Winters and Riley (1992) developed a numerical code to solve the Taylor–Goldstein equation in finite difference form. Using the same second-order difference operators we have modified the code to include the topographic term present in our extended, long-wave version of this equation (2.12). Given the fitted monthly mean N^2 profiles (Fig. 6) and the monthly mean veloc-

ities (Fig. 7) specified at 15 levels, we have calculated the vertical modes using 15 interior grid points. This number is sufficient to resolve the first and second vertical modes as verified by our tests using analytical profiles with known solutions. Since the horizontal velocity profiles have at most two zero crossings, the first two vertical modes appear to be the ones of greatest importance for the hydraulics of the exchange flow. In the following description we continue to use the convention that $j = +n$ and $j = -n$ denote waves belonging to mode n whose speeds in the absence of a background velocity obey $c_n = -c_{-n}$.

Figure 9 shows the propagation speeds for $j = \pm 1$ and $j = \pm 2$ modes for each month. The speeds are represented by arrows which point to the right (left) if the wave propagates towards the Red Sea (Gulf of

Hanish Sill (November 1995 - November 1996)

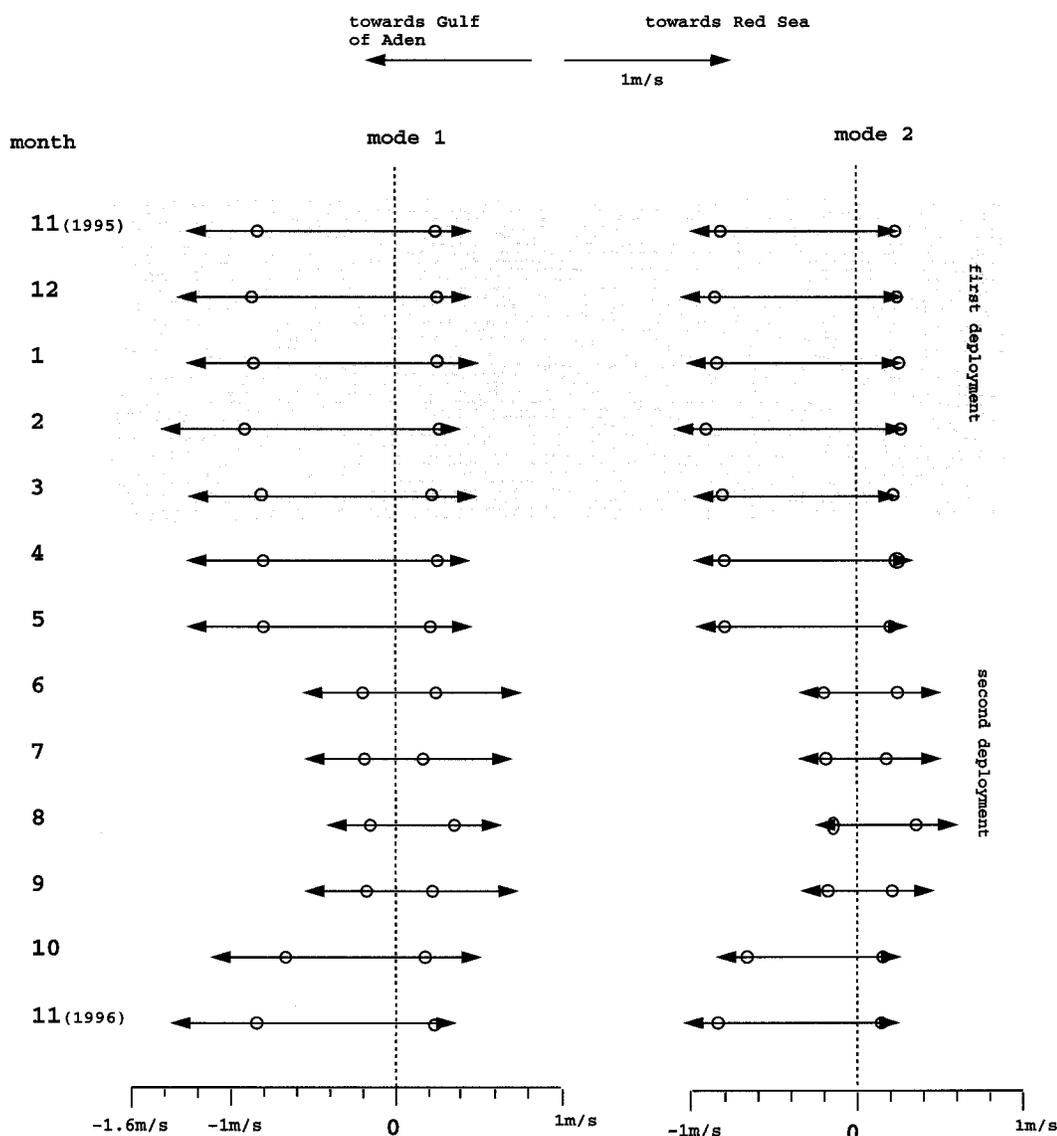


FIG. 9. The arrows indicate phase speeds c_j for the first and second internal modes ($j = \pm 1, \pm 2$) based on the monthly mean sill velocity and stratification. For each month listed the four arrows give the speed and direction of modes $j = \pm 1$ (left column) and $j = \pm 2$ (right column, in m s^{-1}). The circles indicate the extremes U_{\min} and U_{\max} of the velocity profile. The results in the unshaded region include second deployment measurements in which full stratification was known. The shaded region covers first deployment measurements in which the velocity but not the stratification was known. The results were obtained using stratification from Apr and Nov 1996 (4/96 and 11/96). Uncertainty estimates based on varying the stratification and velocity profiles (as described in appendix C) and on estimates of numerical error are less than 8 cm s^{-1} in each case.

Aden). For a given mode, a pair of arrows pointing in opposite (the same) directions indicates subcritical (supercritical) flow with respect to that mode. If one or both of the arrows has zero length, the flow is critical with respect to that mode. The two circles on each set of arrows represent the extremes (U_{\min} and U_{\max}) of the mean velocity profile for that month. Speed uncertain-

ties are within $\pm 8 \text{ cm s}^{-1}$ for each month. Our estimates of uncertainty are due to extrapolation of the velocity profiles, fitting of the density profiles, the presence of a residual barotropic flow in the total transport, and expected numerical errors (see appendix C for full discussion).

We first draw the reader's attention to the results based

on second deployment (months 4–11 in 1996) data, which are shown in the unshaded part of Fig. 9. In all second deployment months the flow is subcritical with respect to both modes, although there are some instances when flow is moderately close to a critical state with respect to one or both modes. This can be seen during months 4, 5, 10, and 11; c_1 and c_2 (the speeds of the two waves propagating into the Red Sea) are considerably reduced in relation to their partners c_{-1} and c_{-2} . For example, in month 11 c_{-1} is -1.37 m s^{-1} and c_{+1} is 0.34 m s^{-1} so that signal propagation in the direction of the Red Sea is relatively slow. For the second mode, the speeds are -1.04 and 0.24 m s^{-1} , respectively, leading to a similar interpretation. During the summer months (6–9) there is an imbalance favoring the wave propagating *toward* the Red Sea. For example, month 8 has speeds ($c_{-2} = -0.34$, $c_2 = 0.58$) indicating relatively slow propagation towards the Gulf of Aden. As noted by PJMK and explained in the introduction, it is somewhat surprising to find subcritical flow at the sill as it appears to be the most likely point of hydraulic control in the BAM.

For completeness we have recomputed $c_{\pm 1}$ and $c_{\pm 2}$ from the first-deployment months (Nov 1995–Mar 1996) using our continuous model. This calculation incorporates direct, monthly mean ADCP data along with the stratification from April 1996 and November 1996. (Complete stratification during the first deployment was unavailable.) The results are shown in the shaded region of Fig. 9. Again, we see that the flow is subcritical with respect to both modes, but that the speeds of the waves attempting to enter the Red Sea are small, possibly indicating a nearby control section. These results imply a slightly subcritical flow, whereas PJMK indicate that the flow is right at the critical point with respect to mode 2.

The modes for each month display a common behavior in their relationship to the measured velocity profile. The phase speeds always lie outside, in some cases barely outside, the window of the velocity extremes. This can be seen in Fig. 9, in both the shaded and unshaded regions, as each arrow stretches beyond the circles representing the background velocity extremes. Since the phase speeds are never inside that window, there are no critical levels with respect to these modes.

In summary, the flow is subcritical with respect to each mode, though barely so in the winter months. The phase speeds based on monthly mean data lie outside the extremes of monthly mean velocity for each month, therefore critical levels do not arise. The mismatch in the magnitude of the negative and positive extremes of background velocity lead to similar mismatches in the propagation speeds for the two waves belonging to a particular mode. For the second mode, the phase speeds for most months lie very close to these extremes, suggesting incipient critical level behavior. This behavior is reflected in the positions of the zero crossings of $\tilde{w}(z)$ (not shown), which lie close to the levels of the velocity extremes.

6. Criticality with respect to instantaneous flows

The sill flow in the BAM is strongly influenced by diurnal and semidiurnal tides and by subtidal disturbances. The subtidal disturbances typically have periods of a few days to a few weeks and are forced in part by wind stress variations over the strait (Seidler 1968). The detailed nature of the fluctuations are the subject of a separate investigation on the new Bab al Mandab data; here we simply note that they can be comparable in strength to the mean flows in the strait and at times can cause the top-to-bottom flow to be unidirectional. All of these features are apparent in the Fig. 10 time histories of ADCP velocity at three depths over months August–November 1996. We have made several phase speed calculations based on two-hour mean velocity profiles and instantaneous stratifications observed at various times during this period. These calculations are presented in order to show how dramatically the criticality of the instantaneous flow can depart from the monthly mean. A comprehensive study of time dependence is beyond the scope of this work.

We have computed phase speeds for ten realizations of the 2-h mean flow, corresponding to the daily extreme of the tidal ebb and flood on five selected days. These days are indicated on the Figs. 10a and 10b time series. The 2-h mean velocity profiles from two of these days are displayed in Figs. 11a,b. Figure 11a shows the results for 10 August 1996, which falls during a period of average tides and little subtidal activity. Even in the presence of such moderate tides, the maximum tidal flood (Red Sea inflow), which occurs at hour 17, is strong enough to reverse the normal deep and shallow outflows for August, creating inflow over the entire water column. Figure 11b shows the same information for 18 November, which falls during a period of strong outflow at 20-m depth due to a subtidal disturbance (see the Fig. 10b time record, in which 18 November is marked). In this case, top-to-bottom outflow occurs throughout the entire day.

The Richardson number profiles for the maximum flood and ebb on 18 November (hours 13 and 19) are shown in Fig. 12. Excluding values near the top and bottom (which are subject to large uncertainty due to extrapolation of the velocity), the lowest values range from 0.20 to 0.46 at maximum ebb, and from 0.08 to 0.29 at maximum flood. (The lowest value found for the monthly mean flow for this month was 0.25.) These values occur at depths from 30 m to 60 m. (Depths without marked values have Richardson numbers that are large enough to be off the horizontal scale of the figure.)

Figure 13 shows the phase speeds for instantaneous flows on the five selected days. For each day two results are shown corresponding to the maximum tidal flood (top set of arrows) and ebb (bottom set of arrows). Apart from 10 August critical and/or supercritical flows occur during the maximum ebb tide on all four days (the lower

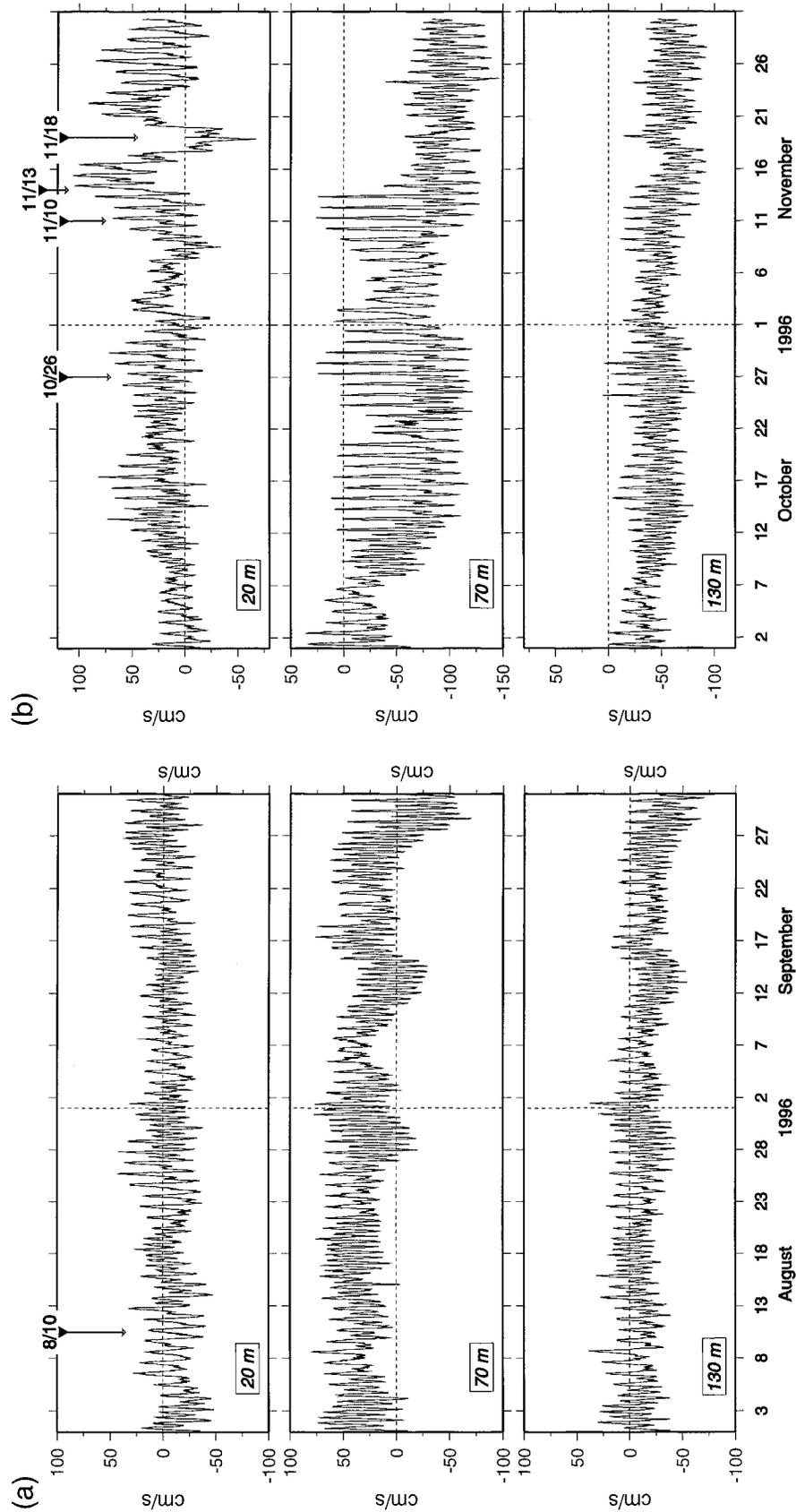


FIG. 10. Time histories of ADCP velocity at three depths over the period (a) Aug–Sep 1996 and (b) Oct–Nov 1996. Arrows indicate the days for which further analysis was carried out.

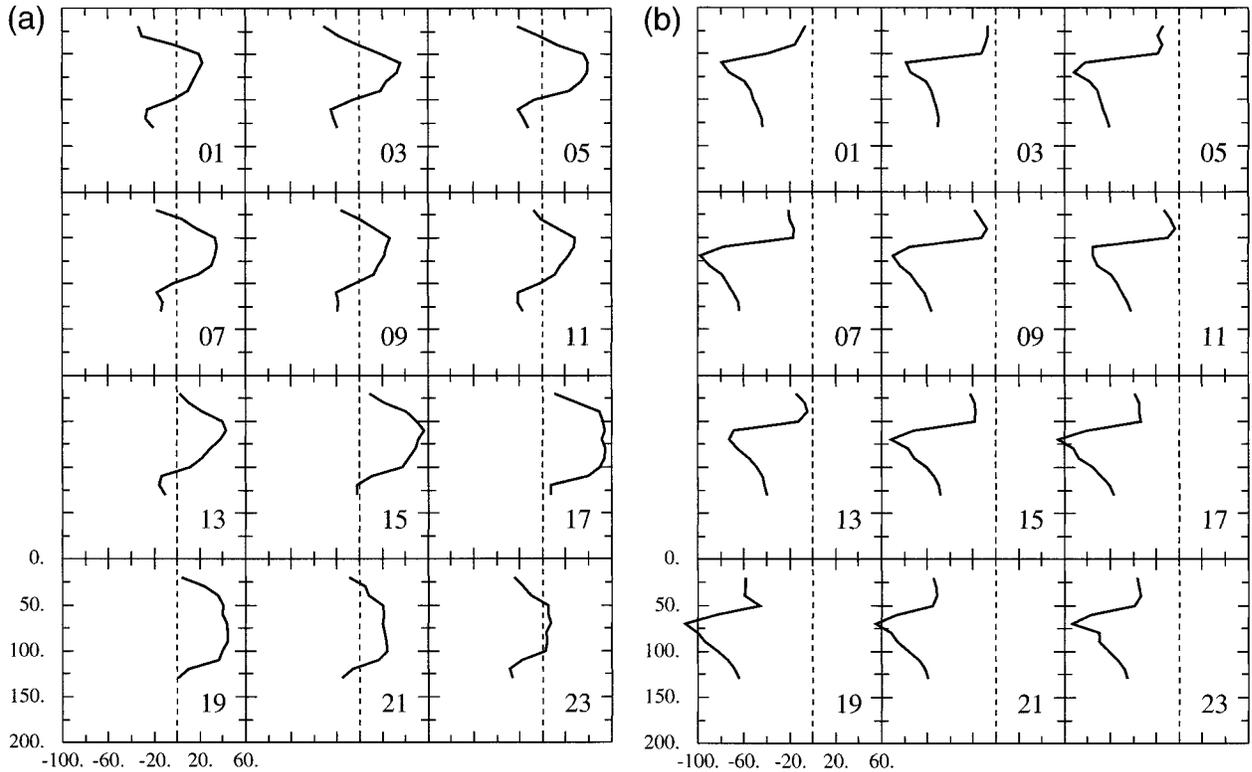


FIG. 11. Two-hour mean ADCP velocity profiles in centimeters per second on (a) 10 Aug 1996 and (b) 18 Nov 1996. The maximum flood (into the Red Sea) and ebb occur on 10 Aug at hours 17 and 23 respectively. On 18 Nov the maximum flood and ebb occur at hours 13 and 19.

set of arrows). The circles in Figure 13 indicate the extremes of the 2-h mean velocity and it can be seen that in all cases but one (mode 2 on 13 November), the phase speeds continue to lie just outside U_{\min} and U_{\max} . Critical levels are thereby avoided, as was the case for the monthly means. These results indicate that the tidal and subtidal motions easily push the flow into a critical or supercritical regime with respect to both modes. The criticality of the flow is evidently a strongly intermittent feature.

Estimation of the average propagation speed based on the characteristic depth and velocity along the BAM axis yield values on the order of 1 m s^{-1} for a first or second baroclinic mode. The corresponding average travel distance over the 12-h tidal period is $\cong 40 \text{ km}$, which is only a fraction of the 180 km length of the BAM. A signal propagating from the Gulf of Aden into the Red Sea would therefore spend many tidal cycles in the strait. The results of this section suggest that the direction of propagation can be temporarily, and perhaps permanently, reversed by intermittent supercritical conditions. The implications for hydraulic control are discussed further below.

7. Discussion

By developing an extended version of the Taylor–Goldstein equation for long waves [see Eq. (2.12)], we

have gained the ability to calculate the topographically modified, internal-gravity modes for given profiles of velocity and density. An application of the equation to given velocity and stratification at a cross section of a particular strait simply requires formulation of the topographic function $T(z) = b^{-1} db/dz$. Our central conclusions concerning the first and second baroclinic modes calculated at the BAM sill follow from an intriguing constraint observed for the long-wave phase speeds. Specifically, we nearly always observe that the speeds c_{-j} and c_j of the two waves belonging to the first ($j = 1$) or second ($j = 2$) vertical mode obey

$$c_{-j} < U_{\min} < U_{\max} < c_j, \quad (7.1)$$

where U_{\min} and U_{\max} are the minimum and maximum of the velocity profile. No unstable long waves are found. An immediate consequence is that none of the waves have critical levels. Further, for the monthly mean profiles, each of which have $U_{\min} < 0 < U_{\max}$, the flow must be subcritical. For instantaneous profiles (7.1) continues to hold, although the velocity profile can be unidirectional, allowing for critical ($c_j = 0$) or even supercritical (c_{-j} and c_j have the same sign) conditions. We conclude that hydraulically critical flow at the sill is an intermittent feature largely at the mercy of tidal and subtidal motions.

Our analysis of idealized profiles suggests that the

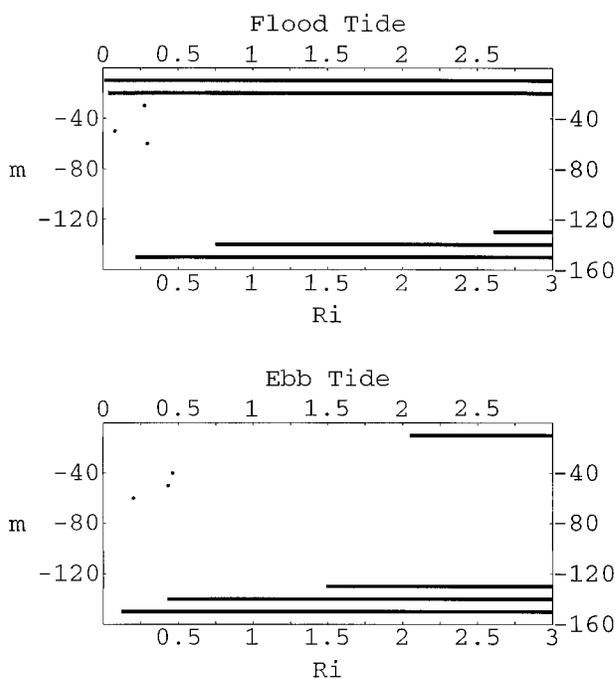


FIG. 12. Richardson number at depth intervals for tidal extremes (hours 13 and 19) 18 Nov 1996. The top frame corresponds to the maximum flood tide (entering the Red Sea) while the bottom frame corresponds to the maximum ebb tide. Uncertainty bars were obtained by varying the velocity and stratification extrapolation near the top and bottom of the water column, as explained in appendix C. The resulting uncertainty bars at depths of 10, 20, 130, 140, and 150 m are quite large. Also, the absence of a R_i value at a particular depth means that the value is >3.0 .

observed frequent proximity of the phase speeds to U_{\min} and U_{\max} may be a characteristic of flows that are marginally stable, perhaps as a result of prior mixing. Richardson numbers calculated from the monthly mean and instantaneous profiles seem to be consistent with this hypothesis. Middepth values of R_i were frequently found to be <1 and sometimes $<1/4$.

When the flow is critical with respect to a particular wave, the vertical structure of that wave gives an indication of the manner in which the flow is being controlled. For example, we would expect that critical control with respect to a wave that is bottom trapped would involve regulation of the deepest part of the flow. PJMK found a nearly critical, second vertical mode that was bottom intensified in terms of the layer transports in their 3-layer model. They understood this structure to indicate the existence of a mode-2 control on the winter transport of Upper Red Sea Deep Water. Unfortunately, comparisons with the vertical structures of the continuous modes found here are difficult. For one thing, our mode-2 waves for the monthly mean winter flow are not critical. For another, the critical mode-2 waves that we have found for instantaneous conditions involve background velocity profiles quite different from the mean velocity profiles that PJMK based their results on. In brief, it is difficult to find a basis for comparison.

This study raises a number of other issues that need to be addressed on a basic level. The most obvious concerns the significance of critical flow in a highly time-dependent setting. If it is indeed the case, as our results suggest, that the criticality of the BAM sill flow is highly intermittent, what does this mean for overall hydraulic control of the exchange flow?⁵ [One of the few modeling efforts to address this problem is Helfrich's (1995) study, but the special nature of the flows involved make it difficult to relate the results to the BAM.] Clearly, the fundamental issues concerning signal propagation and regions of influence must still hold. If the long internal waves of a certain mode are unable to propagate into the Red Sea from the Gulf of Aden, then a certain degree of self-determination (a central feature of hydraulic control) is implied for the Red Sea. As we noted earlier, a first or second baroclinic wave propagating from the Gulf of Aden into the Red Sea would spend at least two or three tidal cycles in the strait. The direction of propagation might even be temporarily reversed by intermittent supercritical conditions, or the information carried by the wave might be lost due to encounters with hydraulic jumps. Clearly, a determination of whether long waves can propagate all the way through the strait in the presence of fluctuating currents is a matter that requires analysis of the flow along the entire strait, not just at the sill.

A second mystery, which is avoided in the present problem but which is of great concern for the critical control of *steady* exchange flows in general, concerns the presence of critical levels. If a steady, continuously stratified exchange flow is hydraulically controlled at some section, then the wave that is stationary there has a critical level z_c at the zero crossing of the velocity profile. Such a wave will generally be unstable. This situation is avoided in two-layer models of exchange flow, where the velocity is always nonzero, and in continuously stratified unidirectional flow (e.g., Armi and Williams 1993) in which stationary waves have no critical levels. Remarkably, long-wave critical layers are also avoided in our BAM observations; stationary waves only arise at the sill when the flow is unidirectional. Nevertheless, we might consider a hypothetical, steady exchange flow with continuously varying velocity and stratification. To what extent do the classical ideas about hydraulic control apply when stationary long waves occur in such a flow?

Some guidance might be gained from the literature on severe downslope winds in the atmosphere (e.g., Pelletier and Clark 1979; Smith 1985) in which both hydraulic behavior and critical levels can arise. Instability and wave breaking about the critical level can result in the formation of a homogeneous wedge of fluid at and

⁵ This is hardly a new issue, having been placed at the top of the list of items for future study at a NATO workshop on sea straits in 1989 (see Pratt 1990).

**Hanish Sill (August 1996 - November 1996)
'Instantaneous' Modes**

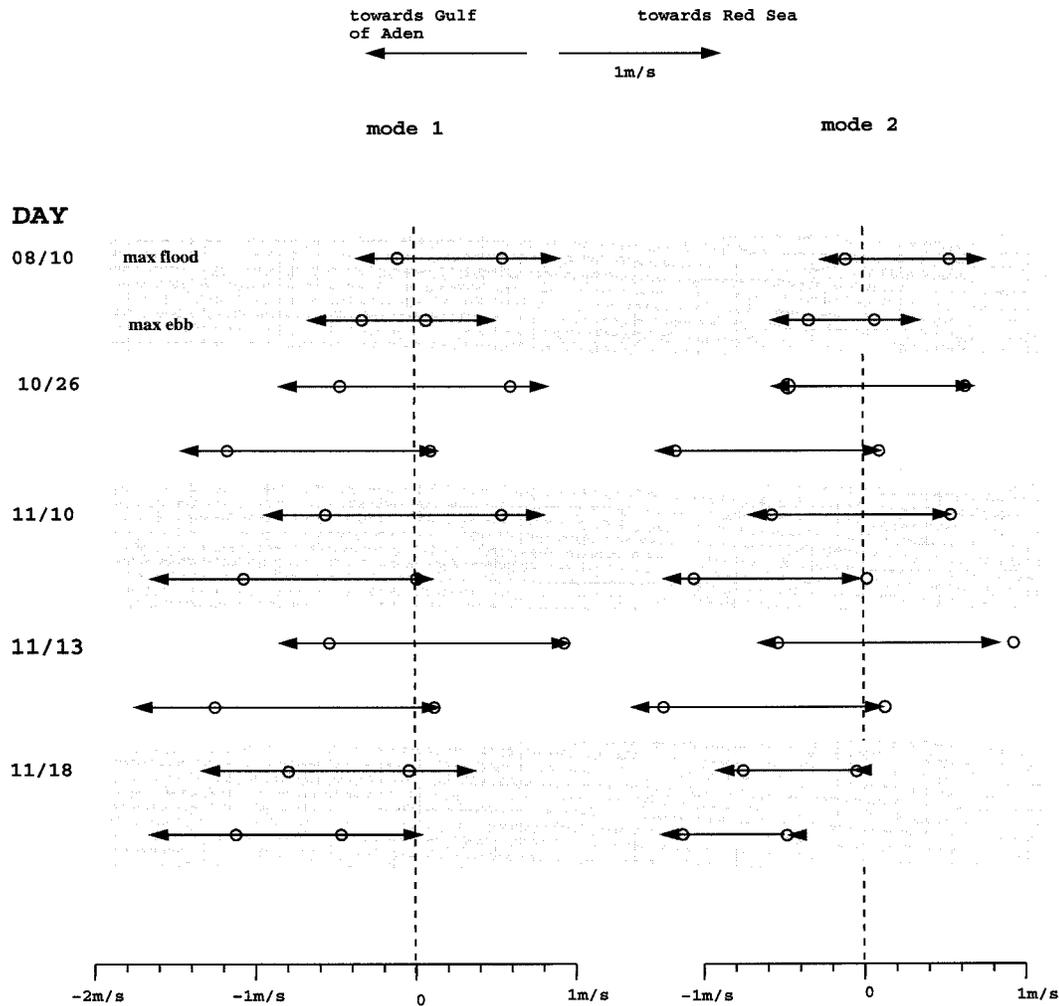


FIG. 13. Phase speeds of first and second modes for tidal extremes on dates 10 Aug; 26 Oct; 10, 13, 18 Nov 1996. As in Fig. 11, arrows represent the phase speeds and circles sit on the extremes (U_{min} and U_{max}) of the velocity profile used to calculate that particular mode. For each day the upper set of arrows corresponds to maximum flood tide and the lower set of arrows corresponds to maximum ebb tide. Uncertainty estimates (as described in appendix C) are less than 8 cm s^{-1} for all calculations. Note that only one calculation (the flood tide, mode 2 result form 13 Nov) shows the presence of a critical level (a phase speed within the range of the velocity). However, the difference between the maximum velocity and the phase speed for this realization is within the uncertainty estimate.

immediately downstream of the mountain crest. The base of the wedge acts as a perfect reflector of wave energy, dynamically insulating the underlying fluid from the flow aloft. This underlying fluid spills down the slope in the manner of a supercritical reduced-gravity layer. Hydraulic supercriticality can be diagnosed by solving the Taylor–Goldstein equation in this layer, subject to the condition for a pliant upper boundary. This upper boundary is simply the base of the homogeneous wedge. Laboratory and numerical demonstrations of the formation of the wedge appear in Wilkinson and Wood (1985) and Bachmeister and Pierrehumbert (1988), re-

spectively. An analytical model for the evolution of the lower layer is discussed by Smith (1985).

In the BAM, the processes just described would presumably lead to a layer of homogeneous fluid about the zero crossing of the velocity profile. However, such homogenization is not observed; the zero crossing of the velocity tends to lie within the continuously stratified intermediate depths of the water column (e.g., Fig. 6 of PJMK), at least where velocity measurements have been made (the sill and narrows). Also, the development of a critical level into a perfectly reflecting boundary is a process that takes time and it is not clear how this de-

velopment would occur in the presence of tides. Nevertheless, there are reasons that one might pursue the meteorological analogy further. First, Richardson numbers $\leq O(1)$ are observed at levels close to the zero crossing of the monthly mean velocity. Such values are consistent with prior mixing though, as noted above, they are not associated with outright homogenization. Second, there is a layer of relatively homogeneous fluid that forms on the downslope south of the narrows, namely the stationary wedge of Gulf of Aden Intermediate Water (GAIW) that exists for most of the year (Smeed 1997).⁶ This wedge bears remarkable similarity to the wedge of homogeneous fluid that forms around the critical level in meteorological simulations (e.g., Bachmeister and Pierrehumbert 1988), laboratory experiments (Wilkinson and Wood 1985), and other field studies (Farmer and Armi 1999). One difficulty with this analogy is that GAIW is colder than both the surface inflow and underlying outflow, so that it cannot have been formed as a result of mixing between the two. Also, estimates of wave speeds based on PJMK's three-layer model suggest that the flow is substantially subcritical just upstream of this wedge.

Acknowledgments. This work was funded by Grants 00014-95-1-0456 (Pratt and Deese), 00014-96-1-1227 (Murray), and 00014-95-1-0025 (Johns) from the Office of Naval Research. The authors are grateful to Rainer

⁶ In the summer this wedge move up into the BAM, forming the intermediate Red Sea inflow.

Zantopp and Eddie Weeks for assistance in processing the BAM data and drafting figures.

APPENDIX A

Finite Difference Approximation of the Topographic Term in (2.12)

Winters and Riley (1992) discuss the second-order finite difference representation of the Taylor–Goldstein equation and its extended version including viscosity and thermal diffusivity. Our (2.12) has been solved numerically by first multiplying that equation by $(U - c)^2$ and then employing same second-order representations used by Winters and Riley. In differencing our new topographic term

$$(U - c) \frac{d}{dz} [(U - c)T\bar{w}]$$

it should be kept in mind that T may be singular at the channel bottom $z = 0$. At this point, the behavior of $T(z)$ will normally be known analytically, either because the entire cross section is idealized (as in our example with trapezoidal topography) or because actual topographies will have to be analytically approximated. To the extent that this analytical representation is trustworthy, it is dangerous to approximate $d[(U - c)T\bar{w}]/dz$ as $[U_{j+1}T_{j+1}w_{j+1} - U_{j-1}T_{j-1}w_{j-1}]/2\Delta z$ (where j is a grid point index on a grid of spacing Δz) since this representation becomes inaccurate where $T(z)$ singular. Instead, the differentiation should first be performed term by term and the result differenced, leading to the representation

$$\frac{d}{dz} [(U - c)T\bar{w}] = \frac{(U_{j+1} - U_{j-1})T_j w_j + (U_j - c)w_j \left(\frac{dT}{dz} \right)_j + (U_j - c)T_j (w_{j+1} - w_{j-1})}{2\Delta z} + O(\Delta z)^2.$$

This form allows the analytical value of dT/dz to be specified directly.

APPENDIX B

Couette Flow over an Obstacle

Consider an inviscid, homogeneous, two-dimensional flow with uniform shear ω over an obstacle of elevation h . In order to remove free surface gravity waves, the fluid is bounded above by a rigid plate of constant elevation z_T . In the long-wave (hydrostatic) limit the vorticity ω is dominated by the term $\partial u/\partial z$ and therefore

$$u(z) = u_0(x) + (z - h(x))\omega,$$

where u_0 is the ground-level velocity. Conservation of volume flux $Q = \int_h^{z_T} u dz$ then requires that

$$u_0 = \frac{Q - \frac{1}{2}(z_T - h)^2\omega}{(z_T - h)}$$

so that there is a unique u_0 , and therefore a unique velocity profile, for each given value of the bottom elevation h . The hydrostatic pressure along the bottom can be computed from the Bernoulli equation:

$$\frac{p_0}{\rho} = B_\infty - \frac{(u_0)^2}{2} - gh,$$

where B_∞ denotes the value of the Bernoulli constant

along the bottom streamline. Since the pressure p_0 is the same at equal values of h on either side of the obstacle, the total form drag by the obstacle is zero and the obstacle can exercise neither hydraulic control nor upstream influence.

APPENDIX C

Uncertainty Estimates

We attempted to estimate uncertainties in the phase speed calculations by varying features in the velocity and stratification we were uncertain about over reasonable ranges. For example, we varied the way in which the ADCP velocity profile, which is known only to within 15 m of the surface and 20 m of the bottom, was extrapolated. In one case we extrapolated linearly to the top and bottom using the uppermost or lowermost known pairs. In another, we linearly extrapolated using half the slope determined from the method just described. Our analytical fit of $\sigma_\theta(z)$ was very accurate for all months except August; after attempting a number of analytical fits, we settled on a piecewise linear fit for August. For months November 1995–March 1996, ADCP data was collected, but corresponding CTD data was not. We therefore utilized fits for $\sigma_\theta(z)$ from November 1996 and April 1996 in analyzing this first deployment data. The uncertainty resulting from any of these variations was typically a few centimeters per second or less.

Another source of uncertainty arose from the presence of a residual barotropic flow in the area-integrated transport over the sill sections. That is, we calculated the area integral of $u(z)$ using the ADCP values of u , assuming them to be y independent and using our idealized topography [Eq. (2.5)]. We then divided the result, a volume transport, by the cross-sectional area to get an average velocity \bar{u} . As explained by PJMK, \bar{u} should be positive and less than a few centimeters per second in order to account for Red Sea annual evaporation and observed sea level changes due to seasonal atmospheric pressure fluctuations. The calculated \bar{u} was, in fact, this small during all months except June and July 1996, where $\bar{u} \cong 6\text{--}7 \text{ cm s}^{-1}$. These values were simply included in the uncertainty bar estimates.

The final source of uncertainty considered was numerical error due to the limited number (15) of vertical grid points. Using the numerical results for the hyperbolic tangent $U(z)$ and triangular bottom profile discussed in section 3 we performed some variable resolution runs with first and second modes having vertical scales comparable to those calculated numerically. These calculations indicate propagation speed errors of 1%–3%.

The total uncertainty estimate for a given mode was obtained simply by adding the magnitudes of the individual uncertainties together. This final value

ranged from 1–8 cm s^{-1} . The same components were considered in calculating uncertainties in the instantaneous modes, with the exception of the residual barotropic flow, which is expected to be large for the tidal extremes and was therefore not included in uncertainty bars.

REFERENCES

- Armi, L., and R. Williams, 1993: The hydraulics of a stratified fluid flowing through a contraction. *J. Fluid Mech.*, **251**, 355–375.
- Bachmeister, J. T., and R. T. Pierrehumbert, 1988: On high-drag states of nonlinear stratified flow over an obstacle. *J. Atmos. Sci.*, **45**, 63–80.
- Baines, P. G., 1995: *Topographic Effects in Stratified Flows*. Cambridge University Press, 482 pp.
- Booker, J. R., and F. P. Bretherton, 1967: The critical layer for internal gravity waves in a shear flow. *J. Fluid Mech.*, **27**, 513–539.
- Bower, A. S., H. D. Hunt, and J. F. Price, 2000: The character and dynamics of the Red Sea and Persian Gulf outflows. *J. Geophys. Res.*, **105**, 6387–6414.
- Case, K. M., 1960: Stability of an idealized atmosphere. I. Discussion of results. *Phys. Fluids*, **3**, 149–154.
- Drazin, P. G., and W. H. Reid, 1981: *Hydrodynamic Instability*. Cambridge University Press, 527 pp.
- Farmer, D., and L. Armi, 1999: Stratified flow over topography: The role of small-scale entrainment and mixing in flow establishment. *Proc. Roy. Soc. London*, **455A**, 3221–3258.
- Fedorov, K. N., and S. L. Mechanov, 1988: Structure and propagation of Red Sea waters in the Gulf of Aden. *Oceanology*, **28**, 279–284.
- Garrett, C., K. Speer, and E. Tragou, 1995: The relationship between water mass formation and the surface buoyancy flux, with application to Phillips' Red Sea model. *J. Phys. Oceanogr.*, **25**, 1696–1705.
- Helfrich, K. R., 1995: Time-dependent two-layer hydraulic exchange flows. *J. Phys. Oceanogr.*, **25**, 359–373.
- Howard, L. N., 1961: Note on a paper of John Miles. *J. Fluid Mech.*, **10**, 509–512.
- Maillard, C., and G. Soliman, 1986: Hydrography of the Red Sea and exchanges with the Indian Ocean in summer. *Oceanol. Acta.*, **9** (3), 249–269.
- Maxworthy, T., 1997: A frictionally and hydraulically constrained model of the convectively-driven mean flow in partially enclosed seas. *Deep-Sea Res.*, **44**, 1339–1354.
- Miles, J. W., 1961: On the stability of heterogeneous shear flows. *J. Fluid Mech.*, **10**, 495–508.
- Murray, S. P., and W. Johns, 1997: Direct observations of seasonal exchange through the Bab al Mandab Strait. *Geophys. Res. Lett.*, **24**, 2557–2560.
- Peltier, W. R., and T. Clark, 1979: The evolution and stability of finite amplitude mountain waves. Part II: Surface wave drag and severe downslope windstorms. *J. Atmos. Sci.*, **36**, 1498–1529.
- Pratt, L. J., 1990: *The Physical Oceanography of Sea Straits*. Kluwer, 587 pp.
- , W. Johns, S. M. Murray, and K. Katsumata, 1999: Hydraulic interpretation of direct velocity measurements in the Bab al Mandab. *J. Phys. Oceanogr.*, **29**, 2769–2784.
- Seidler, G., 1968: Schichtungs- und bewegungsverhältnisse am sudausgang ds roten meeres. "Meteor" *Forschungsergeb.*, **A** (4), 1–76.
- Smeed, D., 1997: Seasonal variation of the flow in the strait of Bab al Mandab. *Oceanol. Acta.*, **20**, 773–781.
- , 2000: Hydraulic control in three-layer exchange flows: Application to the Bab al Mandab. *J. Phys. Oceanogr.*, **30**, 2574–2588.
- Smith, R. B., 1985: On severe downslope winds. *J. Atmos. Sci.*, **42**, 2597–2603.

- Tragou, E., and C. Garrett, 1997: The shallow thermohaline circulation of the Red Sea. *Deep-Sea Res.*, **44**, 1355–1376.
- , —, R. Outerbridge, and C. Gillman, 1999: The heat and freshwater budgets of the Red Sea. *J. Phys. Oceanogr.*, **29**, 2504–2522.
- Wilkinson, D. L., and I. R. Wood, 1985: The formation of an intermediate layer by horizontal convection in a two-layer shear flow. *J. Fluid Mech.*, **136**, 167–187.
- Winters, K. B., and J. J. Riley, 1992: Instability of internal waves near a critical level. *Dyn. Atmos. Oceans*, **16**, 249–278.